Appendix D Resource Variability Parameters

There are three basic resource variability parameters for renewables with variable resources (i.e. wind and solar) that are calculated for each period in ReEDS before the linear program optimization is conducted for that period. These include capacity value, operating reserve, and surplus. For each, a marginal value is calculated, which applies to new installations in the period, and an "old" value is calculated, which applies to all the capacity built in previous periods. This section describes the statistical assumptions and methods used to calculate these values.

These variable-resource parameters are calculated for a source from which the variable-resource renewable energy (VRRE) is generated and a sink to which the energy is supplied. The source is always a supply region. The user must specify the regional level for the sink. It can be a balancing authority (BA), a regional transmission organization (RTO), a NERC region, or an entire interconnect. The "old" values for these variable-resource parameters are calculated for each sink but not for each source since the old value is a single value for all the variable resource supplied to the sink.

D.1 Data inputs for the calculation of resource variability parameters

The inputs required for calculating the resource variability parameters describe the probability distributions associated with loads, conventional generator availability, and VRRE generation. For each, an expected value and standard deviation are calculated.

For loads the expected value, μ_L, is the same as the values used in the "LOAD_PCA" constraint. The standard deviation of the load, μ_L, is found from the load-duration curve of the sink region.

For conventional generator availability, the expected value is the nameplate capacity times 1 minus the forced outage rate.

$$\mu_C = \sum_q CONVCAP_{q,r} \cdot (1 - fo_q)$$

Variance of conventional generator availability is calculated thus:

$$\sigma_C^2 = \sum_q numplants_{q,r} \cdot plantsize_q^2 \cdot fo_q \cdot (1 - fo_q)$$

where

$plantsize_q$ is the input typical size of a generator of type q
$numplants_{q,reg} = CONVCAP_{q,r}/plantsize_q$

The probability distribution associated with conventional generator availability is complicated by the fact that there are can be many conventional generators and each one's availability is a binomial random variable with probability $(1 - fo_q)$ of being one. We largely avoid this complication by first combining the random variables for conventional generator availability, C, with loads, L, in the form of a random variable X where:

$$X = C - L$$

The expected value of X, μ_X, is the sum of the expected values of the other two random variables

$$\mu_X = \mu_C - \mu_L$$

and, since C and L are statistically independent:

$$\sigma_X^2 = \sigma_C^2 + \sigma_L^2$$
$$\sigma_X = \sqrt{\sigma_C^2 + \sigma_L^2}$$

where σ denotes standard deviation and σ^2 is the variance.

We investigated several standard distributions for their ability to fit X. We considered only those distributions that could be bounded below - Beta, Weibull, Gamma, Erlang, Rayleigh, Triangular, Log-logistic, Pareto, Exponential, Uniform, Log-normal, Inverse, Gaussian, and χ^2 To do this we used empirical data from the ERCOT region of Texas and computed three statistics of the goodness-of-fit for this data - χ^2, Anderson-Darling, and Kolmogorov-Smirnov statistics. We based the conventional generation data upon a random sampling of forced generator outages, using empirical outage rate data obtained for ERCOT in 2005. The load data is based upon a Markov chain model we developed from actual empirical data.

Table 25 ranks the distributions as to how well they fit the data relative to each of the three statistics. The beta distribution provides the best fit for the Anderson-Darling statistic and the second best fit for the other two statistics. The beta distribution has the additional advantage that it is bounded both from below and above, similar to the data itself. Figure 9 visually shows how well the PDF of the fitted beta distribution matches the actual data; note that the fit is relatively better at the tails of the distributions which are the areas of greatest interest in our calculations of the resource variability parameters that will be described later.

Table 25: Rankings of Distributions

Ranking		Statistic	
	χ^2	Anderson-Darling	Kolmogorov-Smirnov
1	Triangular	Beta	Weibull
2	Beta	Triangular	Beta
3	Weibull	Gamma	Triangular
4	Gamma	Log-normal	Gamma
5	Log-normal	Uniform	Log-normal
6	Uniform	Exponential	Uniform
7	Exponential	Weibull	Exponential
8	Inverse-Gaussian	Inverse-Gaussian	Inverse-Gaussian

Thus we approximate the combined distribution for $X = C - L$ as a beta distribution with mean μ_X and standard deviation σ_X.

The statistical representation of the output of the VRREs is similar to that of $X = C - L$, although perhaps more complicated due to resource variation, correlations between the VRRE plants and technology change. The standard deviation associated with an individual VRRE site is derived from the hourly data available from the Wind and Solar Integration Study (WSIS) led by NREL (`http://mercator.nrel.gov/wwsi`) The standard deviation of the generation within each ReEDS time slice is easily calculated by standard statistics. To perform a distribution analysis similar to that performed on X, we randomly selected a number of wind sites from the WSIS study and tested how well various standard distributions matched the data. As with X, the beta distribution was a clear winner. Figure 10 shows the power output of, on the left, a single wind site, and, on the right, the combined output of eight randomly selected sites. Other random site selections produced similar charts. Both charts have been fitted with beta probability distribution functions.

Future improvements in the performance of wind and solar technologies are captured in ReEDS through increased capacity factors. These improved capacity factors translate directly

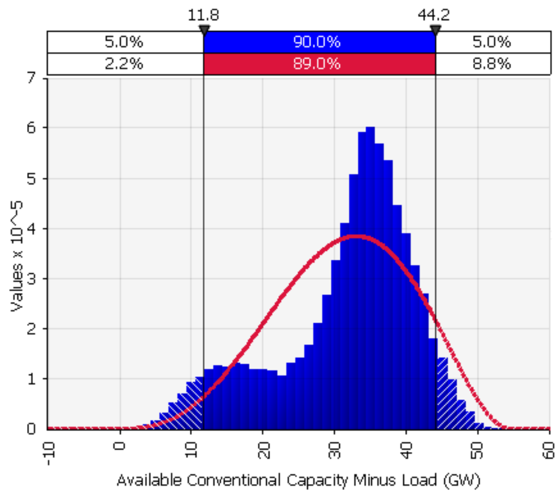

Figure 9: Actual conventional capacity less load and fitted beta PDF

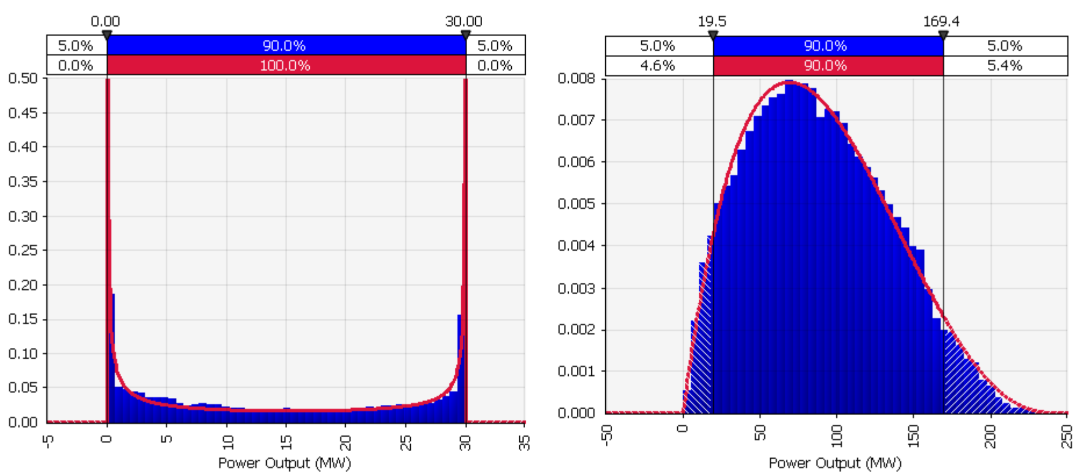

Figure 10: Wind farm power output from a single site (left) and combined output from eight sites (right) with fitted beta PDFs

into improvements in the mean of a VRRE plant's generation output. ReEDS also estimates a new standard deviation for a VRRE plant based on regressions that estimate the new standard deviation as a function of the old standard deviation and the new capacity factor.

In the variable-resource parameters described below the input distributions must represent the generation from all VRRE plants contributing to a sink region, not simply a single plant. The mean value μ_R is easily calculated as the sum of the mean values of the output of the individual contributing VRRE plants. The standard deviation is complicated by the fact that the outputs of the VRRE plants are correlated with one another. For each ReEDS time slice, we have used the WSIS data to develop a correlation matrix (P_{kl}) of the Pearson correlation between each possible pair k,l of region, class, and VRRE, e.g. the peak-time slice correlation matrix has an entry for the correlation between peak-time slice generation from class 5 wind in region 3 and class 2 PV generation in region 14. This P_{kl} matrix is an input to ReEDS. The variance of the VRRE arriving at a sink region r ($\sigma^2_{R_r}$) is then calculated from this correlation matrix P_{kl} through the standard statistical formula:

$$\sigma^2_{R_r} = \sum_{k \in R_r} \sum_{l \in R_r} P_{kl} \cdot \sigma_k \cdot \sigma_l$$

where

R_r is the set of VRRE's contributing to region r

Armed with the mean and standard deviation of all VRRE contributing to a region r, we now have to assume a distribution for the VRRE generation. As with the conventional generation and load, we again used the beta distribution to approximate the VRRE generation with mean μ_R and standard deviation σ_R. This selection of the beta distribution was based on the facts that the beta is a two-parameter distribution with support bounded by a minimum and maximum level, i.e. for VRRE generation the minimum is zero and the maximum is the total nameplate VRRE capacity.

With the probability distributions of the VRRE and of conventional output minus load determined, we can now calculate the variable-resource parameters - capacity value, operating reserve, and surplus.

D.2 Capacity Value

This is the capacity credit given to the VRRE contribution to meeting the reserve margin constraint in each sink region. It is a function of the amount and type of VRREs consumed in the sink region, the dispersion of the VRRE plants contributing the energy, the electric load in the sink region, the variability of the load and the amount and reliability of conventional capacity contributing to the load in the sink region. Generally, as more VRREs are used by the sink region, their capacity value decreases. And as more renewable energy from a particular source is used, the marginal capacity value from that source decreases.

$\mathbf{CVold_r}$ For the total VRRE generation that is to be consumed in sink region r, the capacity credit, $CVold_r$, is the amount of load that can be added in every hour without changing the system reliability in sink region r, i.e. without changing the loss-of-load probability. This added load is the effective load-carrying capability associated with the VRRE contributed to the sink region. Generally, utilities desire to keep this loss-of-load probability close to the equivalent of one day in 10 years or 2.4 hours per year, or per 8760 hours. This equates to a loss-of-load probability of $2.74x10^{-4} = 2.4/8760$.

The first step in estimating CVold is to determine when the maximum allowable loss of load probability is reached without any VRRE in the system. To do this we use the random variable X defined above as the sum of all the conventional generation capacity C available to the sink region r, minus the load L in the sink region r.

$$X = C - L$$

Without any VRRE in the system, the maximum allowable loss of load probability is reached when $B(x) = 2.74x10^{-4}$, where $B(x)$ is the cumulative beta distribution function for the random variable X.

The second step in finding CVold is to find a similar value when VRRE is included in the system. To do this we define a new random variable Y as follows:

$$Y = X + R = C - L + R$$

where

R_r is the random variable representing the VRRE contribution to sink region r.

As before, the maximum allowable loss of load probability is reached when $F(y) = 2.74x10-4$, where $F(y)$ is the cumulative distribution function for the random variable Y. $F(y)$ can be determined by convolving the distributions for X and R. Once Y is determined, the capacity value associated with R is simply Y - X, and the capacity value of the VRRE expressed as a fraction of the total VRRE nameplate capacity, R_r, is

$$CVold_r = (Y - X)/R_r$$

Since convolving two random variables can be a computer intensive calculation and because it must be done many times in a single ReEDS optimization, convolutions are done outside of ReEDS for a range of points with the results saved in tabular format as a function of five parameters—VRRE nameplate capacity divided by peak load, conventional nameplate capacity divided by peak load, VRRE capacity factor, standard deviation of all VRRE generation delivered to sink region r, and the standard deviation of conventional generation minus load $(C - L)$. During a ReEDS run this five-dimensional table is accessed by linear interpolation with these 5 independent inputs to find the capacity value of the VRRE capacity contributing to sink region r, $CVold_r$.

$\mathbf{CVmar_{c,i,r}}$ is the marginal capacity value associated with the addition of class c VRRE capacity in a source region i delivered to a sink region r, is simply the difference in *CVold* before and after the marginal VRRE addition, $\Delta R_{c,i,r}$.

$$CVmar_{c,i,r} = (CVold_r(R_r) - CVold_r(R_r + \Delta R_{c,i,r}))/\Delta R_{c,i,r}$$

D.3 Operating Reserve Requirement

Operating reserve includes spinning reserve, quick-start capability, and interruptible load that can be dispatched to meet unanticipated changes in loads and/or power availability. There is no standard approach for estimating the level of operating reserve required. Some NERC regions assume that operating reserve must be at least as large as the largest single system contingency, e.g. the failure of a nuclear power plant. Others have reasoned that a system should have enough operating reserve to meet 7% of peak load (reduced if hydro is available). We assume in ReEDS that the normal operating reserve ($NOR_{r,m}$) required by a sink region r is proportional to the load ($L_{r,m}$) and conventional generation ($G_{r,m}$) in the region.

VRREs can induce a need for additional operating reserve beyond the usual requirement. ReEDS calculates the total operating reserves induced by all load, conventional generation, and VRREs in the system ($TOR_{r,m}$) and the operating reserves induced at the margin ($ORmar_{r,m}$) by the addition of an increment of VRRE capacity.

TOR$_{r,m}$ is the total operating reserve required in region r due to load, conventional generation, and all existing VRRE capacity contributing to sink region r (R_r). By assuming that the normal operating reserve is a 2-sigma reserve, we can estimate the sigma, $\sigma_{NOR_{r,m}}$, associated with the normal system operation (operating reserve required for load and conventional generation) as:

$$NOR_{r,m} = \frac{0.03 \cdot (L_{r,m} + G_{r,m})}{2 \cdot L_{r,m}}$$

$$\sigma_{NOR_{r,m}} = NOR_{r,m} \cdot (L_{r,m} - R_r)$$

Since the normal system issues that require the normal operating reserve occur independently of the resource variability of VRREs, the variances of the two can be added to give the variance of the total. The total operating reserve is then assumed to be twice the standard deviation of the total.

$$TOR_{r,m} = 2 \cdot \sqrt{\sigma^2_{NOR_{r,m}} + \sigma^2_{R_r}}$$

where

σ_{R_r} is assumed to be the standard variation of the output of all existing VRREs contributing to sink region r.

ORmar$_{c,i,r}$ is the marginal operating reserve requirement induced by the next MW of class c VRRE installed in region i that contributes generation to sink region r. It is calculated as the difference in the operating reserve required with an increment $\Delta R_{c,i,r}$ of additional VRRE capacity, minus that required with only the existing VRRE with the difference divided by the incremental VRRE capacity $\Delta R_{c,i,r}$.

$$ORmar_{c,i,r,m} = \frac{2}{\Delta R_{c,i,r}} \cdot \left(\sqrt{\sigma^2_{NOR_{r,m}} + \sigma^2_{R_r + \Delta R_{c,i,r}}} - \sqrt{\sigma^2_{NOR_{r,m}} + \sigma^2_{R_r}} \right)$$

D.4 Surplus

At high levels of VRRE penetration, there are times when the VRRE generation exceeds that which can be used in the system. This “surplus” VRRE generation must then be curtailed. ReEDS calculates the fraction of VRRE generation from existing VRRE plants ($Surplusold_r$) that is surplus as well as the fraction of generation from new VRRE plants ($Surplusmar_r$) that is surplus. ReEDS uses these surplus values to reduce the useful energy contributed by VRREs, making them less coss-effective generators.

SurplusOld$_r$ is the expected fraction of generation from all the VRREs consumed in sink region r that cannot be productively used, because the load is not large enough to absorb both the VRRE generation and the must-run generation from existing conventional sources. This situation occurs most frequently in the middle of the night when loads are small, base-load conventional plants are running at their minimum levels, and the wind is blowing.

To calculate $Surplusold_r$, we use the random variable Y defined in the capacity value discussion above as the must-run conventional base-load generation M minus the load L plus the VRRE generation R.

$$Y = M - L + R$$

Next, we define the surplus VRRE at any point in time, S, as

$$\text{If } Y < 0, S = 0$$
$$\text{If } Y > 0, S = Y$$

Then the expected surplus μ_S can be calculated from the density function of Y, g(y) as follows:

$$\mu_S = \int_{-\infty}^{\infty} sf(s)ds$$
$$\mu_S = \int_{-\infty}^{0} sf(s)ds + \int_{0}^{\infty} sf(s)ds$$
$$\mu_S = 0 + \int_{0}^{\infty} yg(y)dy$$

The density function of y can be found by convolving the density function of $M - L$ together with the density function of the VRREs, similar to that which was done for the calculation of the VRRE capacity value above. However we found that the expected value of the surplus can be well approximated assuming normal distributions for both $M - L$ and R. With the normal distribution assumption, the value of μ_S can be quickly found in ReEDS with the analytical formula derived below:

Now if we assume, as we did in the *CVmar* and *ORmar* calculations above, that by the central limit theorem, Y can be well approximated by a normal distribution, and we define the standard normal variable Y' as $Y' = (Y - \mu_Y)/\sigma_Y$, then

$$Y = Y' \cdot \sigma_Y + \mu_Y, \text{ and}$$
$$dY = \sigma_Y dY'$$

Thus

$$\mu_S = \int_{0}^{\infty} yg(y)dy$$
$$\mu_S = \int_{-\mu_Y/\sigma_Y}^{\infty} (y'\sigma_Y + \mu_Y) \cdot g(y'\sigma_Y + \mu_Y) \cdot \sigma_Y dy'$$
$$\mu_S = \int_{-\mu_Y/\sigma_Y}^{\infty} \sigma_Y^2 \cdot y' \cdot g(y'\sigma_Y + \mu_Y)dy' + \int_{-\mu_Y/\sigma_Y}^{\infty} \mu_Y \cdot \sigma_Y \cdot g(y'\sigma_Y + \mu_Y)dy'$$

Assuming Y is normally distributed, as stated above:

$$\mu_S = \int_{-\mu_Y/\sigma_Y}^{\infty} \sigma_Y^2 \cdot y' \left(\frac{1}{\sigma_Y\sqrt{2\pi}}\right) \exp\left(\frac{(-y'\sigma_Y + \mu_Y - \mu_y)^2}{2\sigma_Y^2}\right) dy'$$
$$+ \int_{-\mu_Y/\sigma_Y}^{\infty} \mu_Y \cdot \sigma_Y \left(\frac{1}{\sigma_Y\sqrt{2\pi}}\right) \exp\left(\frac{(-y'\sigma_Y + \mu_Y - \mu_y)^2}{2\sigma_Y^2}\right) dy'$$
$$\mu_S = \int_{-\mu_Y/\sigma_Y}^{\infty} \frac{\sigma_Y \cdot y'}{\sqrt{2\pi}} \exp\left(\frac{-y'^2}{2}\right) dy' + \int_{-\mu_Y/\sigma_Y}^{\infty} \frac{\mu_Y}{\sqrt{2\pi}} \exp\left(\frac{-y'^2}{2}\right) dy'$$
$$\mu_S = \frac{\sigma_Y}{\sqrt{2\pi}} \exp\left(\frac{-\mu_Y^2}{2\sigma_Y^2}\right) + \mu_Y\Big(1 - N_{0,1}(-\mu_Y/\sigma_Y)\Big)$$

Where $N_{0,1}$ is the standard normal distribution with mean 0 and standard deviation 1.

Then *Surplusold*$_r$ is the difference between the expected surplus with VRRE, μ_S and the expected surplus were there no VRRE generation consumed in sink region r, μ_{SN}, divided by the total VRRE capacity contributing to sink region r, R_r. Or

$$Surplusoldr = (\mu_S - \mu_{SN})/R_r$$

Normally μ_{SN} would be zero, as the conventional must-run units would not be constructed in excess of the minimum load. However, with our assumption of a normal distribution for Y, we do introduce some non-zero probability that Y could be positive even if there were no VRREs, i.e. that the generation from must-run units could exceed load. Thus, it is important to calculate μ_{SN} and to subtract it from μ_S to remove the bulk of the error introduced by the normal distribution assumption. μ_{SN} is calculated in exactly the same way as μ_S, but with no VRREs included.

Must-run conventional capacity is defined as existing available (i.e., not in a forced outage state) coal and nuclear capacity in sink region r times a minimum turn-down fraction, $MTDF$. The expected value of the must-run capacity of type q available at any given point in time, μ_{M_q}, is thus:

$$\mu_{M_q} = CONVCAPq, r * (1 - FOq) * MTDFq$$

where

$CONVCAP_{q,r}$ is the existing conventional capacity in sink region r of type q.
$MTDF_q$ is 0.45 for old (pre-2006) coal plants,
0.35 for new (post-2006) coal plants,
1.0 for nuclear plants.

SurplusMar$_{\mathbf{c,i,r}}$ is the fraction of generation from a small addition $\Delta R_{c,i,r}$ of class c VRRE installed in supply region i destined for sink region r that cannot be productively used because the load is not large enough to absorb both the VRRE generation and the must-run generation from existing conventional sources. It is calculated as:

$$Surplusmar_{c,i,r} = (\mu_{SR+\Delta R_{c,i,r}} - \mu_S)/\Delta R_{c,i,r}$$

Where $\mu_{SR+\Delta R_{c,i,r}}$ is calculated in exactly the same way as μ_S, but with $\Delta R_{c,i,r}$ MW of VRRE added in region i.